

NJIT

New Jersey's Science &
Technology University

THE EDGE IN KNOWLEDGE

CS 280
Programming Language
Concepts

Classifying Characters

cctype

- standard header file in C++ distribution
- useful utilities for character classification

Testing letters

- Does a particular variable contain a letter?

```
char ch;  
if( ch == 'a' || ch == 'b' || ch == 'c' || ...  
    // that works... but it's hard to type,  
    // annoying to read, and error prone  
  
if( ch >= 'a' && ch <= 'z' )  
    // that's better... as long as letters are  
    // contiguous from a-z (true for ascii)
```

cctype: character classification

```
#include <cctype>

char ch;

if( islower(ch) ) {...} // that's much better!

• islower(ch)    // true if lowercase
• isupper(ch)    // true if uppercase
• isdigit(ch)    // true if a digit
• isalpha(ch)    // true if a letter
• isalnum(ch)    // true if letter or num
• isspace(ch)    // true if whitespace
```

Converting case

- Converting from lower case to upper case:

```
// this is correct... and hard to understand
char upperCh;
if( ch >= 'a' && ch <= 'z' )
    upperCh = 'A' + ch - 'a';
else
    upperCh = ch;

// using something from cctype is better:
char upperCh = toupper(ch);

// Note there is also a tolower()
```

How is ctype implemented?

- Put the test/conversion code into a function
- Better: a table lookup
 - Implementation is an array lookup: use the character as the index into a predefined array, where the contents of the array entry for each character is the classification of that character
- What is the efficiency of this?
 - The table lookup is an example of the classic “space/time” tradeoff in CS: more space, less time
 - Calling a function every time incurs a cost; we can avoid this with inline functions or preprocessor macros